Biodiversity Data Journal 13: e157274
doi: 10.3897/BDJ.13.e157274

Research Article

Navigating uncertainty in museum workflows: genomic data mining and curation of the Diptera collections hosted at RMCA

Lore Esselens[‡], Pia Addison[§], Jacqueline Bakengesa[|,¶], Luis Bota[#,¤,«], Laura Canhanga[¤,«], Domingos Cugala[¤,«], Beatriz Daniel[¤,«], Marc De Meyer[‡], Hélène Delatte[»], Jean-Marc Herpers[^], Kurt Jordaens[‡], Sija Kabota[|], Abdul Kudra[|], Ramadhani Majubwa[|], Aruna Manrakhan[˅], Mirene Mussumbe[¤,«], Maulid Mwatawala[|], Franck Theeten[‡], Didier Van den Spiegel[‡], Sam Vanbergen[‡,¦], Carl Vangestel[^,?], Massimiliano Virgilio[‡]

‡ Royal Museum for Central Africa, Tervuren, Belgium
§ University of Stellenbosch, Stellenbosch, South Africa
| Sokoine University of Agriculture, Morogoro, Tanzania
¶ The University of Dodoma, Dodoma, Tanzania
# National Fruit Fly Laboratory, Chimoio, Mozambique
¤ Eduardo Mondlane University, Maputo, Mozambique
« Centre of Excellence in Agri-Food Systems and Nutrition, Maputo, Mozambique
» Centre de Coopération Internationale en Recherche Agronomique pour le Développement, La Réunion, France
^ Royal Belgian Institute of Natural Sciences, Brussels, Belgium
˅ Citrus Research International, Nelspruit, South Africa
¦ University of Leuven, Leuven, Belgium
? University of Ghent, Ghent, Belgium

Corresponding author: Lore Esselens (lore.esselens@africamuseum.be)

Academic editor: Paolo Biella

Received: 28 Apr 2025 | Accepted: 09 Jul 2025 | Published: 12 Aug 2025

Citation: Esselens L, Addison P, Bakengesa J, Bota L, Canhanga L, Cugala D, Daniel B, De Meyer M, Delatte H, Herpers J-M, Jordaens K, Kabota S, Kudra A, Majubwa R, Manrakhan A, Mussumbe M, Mwatawala M, Theeten F, Van den Spiegel D, Vanbergen S, Vangestel C, Virgilio M (2025) Navigating uncertainty in museum workflows: genomic data mining and curation of the Diptera collections hosted at RMCA. Biodiversity Data Journal 13: e157274. https://doi.org/10.3897/BDJ.13.e157274

Abstract

As part of its extensive Diptera holdings, the Royal Museum for Central Africa (RMCA) houses over 100,000 specimens of Tephritidae and Syrphidae, which represent a critical resource for taxonomic and systematic research. Here, we present a feasibility study evaluating streamlined workflows for genomic data mining and archiving in museum

© Esselens L et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

collections. We analysed DNA yield, quality and sequencing performance from more than 1,400 insect vouchers and found few predictable trends, reflecting the nature of heterogeneous and skewed groups of samples collected under largely unknown field conditions. Regardless, our results show that Illumina short read whole genome sequencing can work well even with degraded insect material. In this context, routine short-read sequencing offers a practical first step for genomic data mining, particularly for large collections. It enables us to reserve more complex and resource-intensive methods for the subset of samples that fail initial sequencing (7% of specimens, in our case). As an outcome of this work, RMCA's archiving system has been adapted to integrate genomic data and metadata alongside traditional specimen records. We argue that genomic data should be treated as an integral component of collection management, enhancing scientific value, supporting long term preservation and improving traceability of genetic resources in natural history collections.

Keywords

natural history collections, museomics, whole genome sequencing, true fruit flies, hoverflies

Introduction

Specimens from natural history collections offer unique taxonomic, evolutionary and ecological research opportunities, particularly for rare or extinct taxa or for species which are difficult to collect from the field (Nakahama 2020). In this context, the insect collections of the Royal Museum for Central Africa in Belgium (RMCA), represent historically important repositories of specimens collected during scientific expeditions and research projects in the Afrotropical Region (Sub-Saharan Africa) over the last 150 years. The collections are especially rich in specimens from the Democratic Republic of Congo, Burundi and Rwanda, primarily from the first half of last century, as well as other parts of Sub-Saharan Africa, with increased geographic coverage from the second half of the century onwards. In recent decades, the RMCA's policies regarding sample collection in the Afrotropical Region have undergone significant changes. These shifts were influenced by the growing international debate on decolonising scientific research (Springer Nature 2022), which gained momentum during the Museum's renovation in 2013-2018, as well as by the implementation of the Nagoya Protocol on Access and Benefit-Sharing (ABS, Secretariat Biological Diversity (2011)). The Nagoya Protocol, which entered into force in 2011, provides a legal framework to ensure fair and equitable sharing of benefits derived from genetic resources, thereby promoting ethical collaborations amongst global institutions. While Belgium signed the Protocol in 2011, it only ratified it and became a Party in 2016, which formally marked the beginning of its legal obligations under the treaty. In response, the RMCA has placed increasing emphasis on equitable partnerships and compliance with international agreements, either by adhering directly to the Nagoya Protocol or by strictly following its principles in African countries where it has not yet been implemented.

While substantial efforts have been made to valorise the RMCA collections through the digitisation of specimen images (see Acknowledgements), comparatively little attention has been given to unlocking the potential of the genetic resources embedded within these collections and making them accessible for scientific research. The continuous advances in genomic technologies are rapidly expanding the possibilities for accessing and analysing historical specimens. Methodological approaches to recover genomic data from specimens in natural history collections, once considered beyond the technological capabilities of most institutions (Colella et al. 2020) are becoming increasingly cost and time effective (Ferrari et al. 2023). In this respect, a growing number of dedicated -omics protocols for degraded DNA from museum vouchers have been developed to mine genomic data from museum collections (Guschanski et al. 2013, Knyshov et al. 2019a, Knyshov et al. 2019b, Mullin et al. 2022). A particular challenge arises with suboptimal samples, which, in the context of this study, refer to specimens typically ranging from a few years to several decades in age. In addition to age, the way in which these specimens were preserved, such as storage temperature and preservation medium, also significantly affects DNA quality and downstream analyses (Carter and Walker 1999, Ferrari et al. 2023. The suboptimal samples are neither as high-quality as freshly collected and properly preserved material nor as degraded as truly historical specimens. Researchers working with such material are in a limbo and must decide between standard, cost- and time-efficient laboratory protocols, which may be inadequate for degraded DNA and more complex, time-consuming and often more expensive procedures tailored to highly degraded samples.

Ready-to-use genomic data recovered from natural history collections hold significant potential for both fundamental and applied research. This is especially relevant for our insect holdings, which include extensive collections of Afrotropical Tephritidae (true fruit flies) and Syrphidae (hoverflies). These collections, comprising over 100,000 specimens, are the focus of active taxonomic and phylogenetic research (De Meyer et al. 2016, Delatte et al. 2019, De Meyer et al. 2020a, De Meyer et al. 2020b, De Meyer et al. 2024, Deschepper et al. 2024a, Deschepper et al. 2024b). The samples span a collection period of up to 120 years, with a notable increase in collecting activity over the past 25 years. They include both older specimens and more recently collected material, preserved either as dry-pinned specimens or in absolute ethanol at −20 °C or −80 °C. A subset of these specimens has been whole genome sequenced as part of multiple research projects (see Acknowledgements), either through non-destructive processing of the entire voucher or by processing selected parts, such as legs.

Here, we report the results of a feasibility study that aimed to: (a) develop and implement standardised workflows for Illumina short-read whole genome sequencing (WGS) of Diptera in the RMCA collections; (b) establish a sustainable archiving strategy to enhance the long-term usability of the resulting genomic datasets; and (c) integrate these genomic data into the collection management system of RMCA to ensure linkage of sequence information with collection specimen. We believe this pilot might serve as a useful starting point for comparable natural history collections looking to routinely mine and curate genomic data.

Materials and methods

Taxon sampling

This study refers to a selection of DNA extracts from 1,405 insect vouchers that were collected between 1997 and 2022 in 54 countries across Africa (925 specimens), Europe (367), Asia (83) and America (30). The dataset is strongly skewed towards Tephritidae (1,296 specimens, representing 79 species from the genera *Bactrocera*, *Ceratitis* and *Dacus*) with a smaller representation of Syrphidae (109 specimens from the genera *Eristalinus* and *Melanostoma*) (Fig. 1; Suppl. material 1). This bias reflects the research activities of the RMCA entomology section, which, over the past few decades, have primarily focused on tephritid fruit flies. In contrast, research on Syrphidae is more recent and, in this study, they were included as a secondary test group to assess the broader applicability of our laboratory and sequencing pipeline across different dipteran families. Furthermore, the specimens from the two groups differ in age distribution, with the Tephritidae showing a more skewed pattern as approximately 70% were collected within the last four years, while the Syrphidae mainly originate from a narrower window between 2015 and 2019. While we fully acknowledge the limitations of this non-uniform sampling, we consider it may offer insights into how specimen age influences high throughput sequencing across a range of realistic conditions.

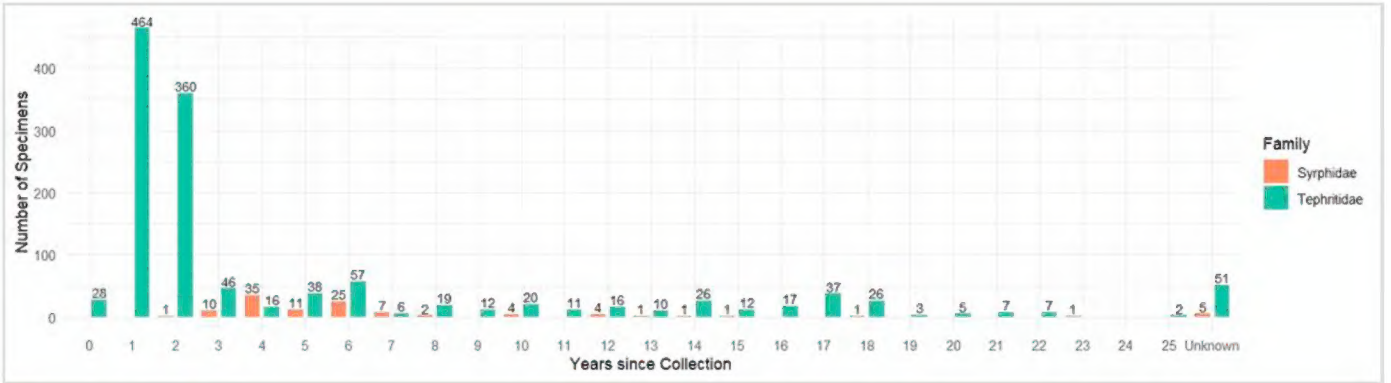

Figure 1. doi

Number of specimens per year since collection (1,405 Tephritidae and Syrphidae processed, including 56 with missing collection year and indicated as “Unknown”).

DNA extraction and whole genome sequencing

The performance of widely used commercial DNA extraction kits (QIAGEN, Suppl. material 2) was preliminarily and semi-quantitatively explored on a subsample of insect vouchers. DNA yields were estimated on three to eight specimens from five species of the target families from seven collection events between 2008 and 2016. This timeframe was selected as deemed to be representative of samples processed in the main experiment (see below). These included three tephritid species, *Zeugodacus cucurbitae*, *Bactrocera dorsalis* and *Dacus bivittatus*, as well as two Syrphidae species, *Eumerus* sp. and *Ischiodon aegyptius*. Specimens were pinned and preserved at room temperature (*I. aegyptius*) or stored in absolute ethanol at -20°C (all other species). Digestion in lysis buffers was implemented on 48 entire specimens or on a single foreleg per specimen to

test if using less material to be less destructive resulted in similar DNA yields (two negative controls with Milli-Q water were also included). The DNA lysates obtained from these samples were divided into four aliquots. Each aliquot was then purified using spin columns from one of the four DNA extraction kits (Suppl. material 2) following the manufacturer's instructions. The concentration of each DNA extract was measured using a Qubit 3 fluorometer (HS DNA Kit, Thermo Fisher Scientific) and the total amount of DNA was inferred from the final elution volume (100 µl). The semi-quantitative comparison of results showed that the kits did not yield major differences in terms of DNA quantity, with the possible exclusion of the MiniElute kit on whole bodies, which had a lower performance (Suppl. material 3). Based on these results, we decided to rely on the kit with the lowest price, i.e. the DNeasy Blood & Tissue Kit (QIAGEN 69506), for the routine processing of specimens.

The suitability of museum Diptera for Illumina short-read WGS was then assessed in a larger experiment involving 1,405 specimens collected between 1997 and 2022. This dataset included taxonomically and temporally diverse specimens, 94.2% being stored in absolute ethanol in freezers at -20°C and -80°C, with the remainder preserved as pinned specimens and dried DNA extracts at room temperature. In all cases, the DNA was extracted using the DNeasy Blood & Tissue Kit (QIAGEN) as per manufacturer's instructions. The DNA extractions were implemented on mostly whole vouchers (99.4%) and, for a very minor part, on insect abdomens or legs. DNA was eluted in volumes ranging from 60-120 µl.

DNA quantity and quality, as well as the HTS performance, were evaluated following DNA extraction. For each sample, the DNA quantity was measured as the total DNA yield per specimen calculated by the DNA concentration, as quantified by using a Qubit 4 fluorometer (HS DNA Kit, Thermo Fisher Scientific) and the elution volume. The DNA quality was measured as the distribution of DNA fragment size by the proportion of the DNA molarity of fragments shorter than 350 bp over the total DNA molarity and DNA purity by the absorbance ratios. The distribution of DNA fragment sizes was quantified on a subset of 902 samples by using a fragment analyser (Advanced Analytical) in combination with the DNF-930 dsDNA Reagent Kit, which covers a range of 75 bp to 200,000 bp (Genomics Core, Leuven, Belgium). The ProSize Data Analysis Software v. 4.0.1.4 (Agilent Technologies) was used to estimate the DNA molarity (nmol/l) of each sample and the proportion of the DNA molarity of fragments shorter than 350 bp over the total DNA molarity. DNA purity was estimated by measuring the absorbance ratios A260/280 and A260/230 with an Implen NanoPhotometer N60 Touch on a subset of 720 samples, using two different absorbance ratios aimed at detecting DNA contamination from different sources such as proteins, phenol, carbohydrates, lipids and salts (*Lucena-Aguilar et al. 2016*). Average absorbances were calculated from three replicated measures to account for measure variability and increased accuracy. Following Lucena-Aguilar et al. (2016), DNA was semi-quantitatively categorised as "pure" ($1.7 < A260/280 < 2.0$ or $1.8 < A260/230 < 2.2$) or "contaminated" (absorbance ratios outside these ranges). The proportion of "pure" DNA samples over the total number of samples was calculated per collection year.

Samples that showed detectable DNA yield (≥ 0.0001 µg) were subjected to Illumina short-read WGS at 10x coverage through the NEBNext Ultra II DNA Library Preparation Kit (New England Biolabs, USA) implemented by Berry Genomics (1,011 DNA samples) or Novogene with standard (178 DNA samples) or low-input library preparation (126 DNA samples), where the genomic DNA was randomly sheared into short fragments, end repaired, A-tailed and ligated with Illumina adapters. Indexing PCR was then performed using Illumina-compatible indexing primers and the libraries were pooled, based on their effective concentrations and expected data output, ensuring optimal sequencing capacity utilisation. The Illumina NovaSeq platform generated 150 bp paired-end reads with a targeted 6 Gb raw data output per sample. To further evaluate the performance of additional steps along the HTS pipeline we then considered: (a) the proportion of high-quality reads (i.e. with a Phred Q-score higher than 30) and (b) the proportion of reads mapped to a reference genome of *Drosophila melanogaste* (GenBank accession no. GCA_029775095). While (a) provided a first estimate of HTS performance, (b) was used to estimate the proportion of target sequences successfully obtained, as the alignment to a closely-related dipteran reference genome excluded most of the non-dipteran contaminant reads. Raw reads were trimmed using the *fastp* tool (*Chen et al. 2018*) and mapped to the *D. melanogaster* reference genome using the *bwa-mem* command from the burrows-wheeler aligner tool (Li and Durbin 2009).

Data analysis

Relationships amongst voucher age, DNA quality and HTS performance were assessed using general(ised) linear models. As sampling was structured differently for the two families, leading to differences in the distribution of samples over time, models were fitted separately for Tephritidae and Syrphidae and the results were compared descriptively to identify patterns. Furthermore, as they represented the vast majority of samples, only specimens preserved in ethanol were included in the analyses. Voucher age was included as the continuous independent variable. Dependent variables included: (a) DNA quantity, (b) proportion of the DNA molarity of short fragments (< 350bp), (c) proportion of pure DNA samples (according to absorbance ratios A260/280 and A260/230), (d) proportion of high quality reads and (e) proportion of reads aligned to the reference genome of *D. melanogaster*. Trends for all variables, except absorbance ratios, were assessed using linear models (LM) with significance tested by F-statistics. Absorbance ratios were modelled using a generalised linear model (GLM) with a binomial error distribution and significance was tested by Chi-square. Variables were initially assessed for normality and homoscedasticity. Model assumptions, including normality and homoscedasticity, were checked using residual diagnostics and Q–Q plots and, where heteroscedasticity was detected, weighted least squares (WLS) approaches were employed. Voucher age was included as fixed effect (*Zuur et al. 2013*). All models were fitted using the *lme4* package (version 1.1-37) in *R* (version 4.4.3). Visualisation of model fits and observed data was performed using the *ggplot2* package (version 3.5.2) in R, employing bar plots of observed proportions combined with regression lines and confidence ribbons derived from the models to effectively illustrate the relationships between voucher age and DNA quality/HTS performance metrics for each family.

Archiving of DNA samples and genomic data

Gen Tegra-DNA 2D barcoded vials with a silica matrix (https://gentegra.com/gentegra-dna-2/) were used for long-term archiving at room temperature of the extracted DNA. A subset of 792 images from 396 vouchers were generated using a focus stacking imaging system (Zerene Stacker software), composed of commercial photographic equipment (Canon DSLR camera and Canon microlens) (*Brecko et al. 2014*). Images from 292 *Dacus* and *Ceratitis* (Diptera, Tephritidae) specimens were also selected for publication on the https://fruitflies.africamuseum.be/ website. As part of a larger effort to digitise the RMCA collections, we relied on the DaRWIN collection management system to archive DNA samples and genomic data (Adam et al. 2019).

Results

The total DNA yield using four commercial DNA extraction kits from QIAGEN on the same specimen ranged from 0.06 µg to 0.15 µg, with an average of 0.65 µg (± 0.53 µg) for whole bodies and from 0.001 µg to 0.022 µg for legs, with an average of 0.006 µg (± 0.006 µg). The use of the DNeasy Blood & Tissue Kit (QIAGEN) on 1,405 insect vouchers produced yields between 0.00 µg and 8.83 µg, with an average of 2.20 µg (± 1.53 µg). Total DNA yield was significantly affected by voucher age in tephritid specimens ($F_{1,1174}$ = 45.09, p < 0.001) (Table 1, Suppl. material 4), with more recently collected specimens yielding higher DNA amounts. In contrast, no significant relationship between voucher age and DNA yield was found in syrphids ($F_{1,85}$ = 0.38, p = 0.54) (Fig. 2A, Table 1, Suppl. material 4).

Table 1.

Table 1: Results of general(ised) linear models testing the effect of voucher age (0–25 years) on six variables. Significance levels: * P < 0.05; ** P < 0.01; *** P < 0.001; n.s. non-significant. Detailed ANCOVA results are provided in Suppl. material 4.

variables	Tephritidae	Syrphidae
total DNA / voucher	***	n.s.
proportion of short DNA fragments (< 350bp)	n.s.	n.s.
absorbance ratio 260/280	***	n.s.
absorbance ratio 260/230	***	n.s.
proportion of quality reads (Q > 30)	***	n.s.
proportion of aligned reads	n.s.	n.s.

The proportion of the DNA molarity of fragments shorter than 350 bp per sample varied between 0.542 and 1 with an average of 0.93 (± 0.059) and 77.9% of samples having a proportion of short DNA fragments higher than 0.90. Voucher age had no significant effect on the proportion of short DNA fragments in either group ($F_{1,738}$ = 0.34, p = 0.56 for Tephritidae; $F_{1,88}$ = 0.24, p = 0.62 for Syrphidae) (Fig. 2B, Table 1, Suppl. material 4).

According to the A260/280 absorbance ratio, 'pure' DNA was present in only 16% of the samples (116 out of 720), while the A260/230 absorbance ratio indicated that 49% of the samples yielded “pure” DNA (351 out of 720). In Tephritidae, the proportion of pure DNA based on A260/280 decreased significantly in more recent vouchers ($LR\ \chi^2 = 53.04$, $df = 4$, $p < 0.001$), while no significant effect was found in Syrphidae ($LR\ \chi^2 = 0.072$, $df = 1$, $p = 0.79$) (Fig. 2C, Table 1, Suppl. material 4). Similar trends were seen for the A260/230 ratio, with purity decreasing over time in Tephritidae ($LR\ \chi^2 = 42.13$, $df = 4$, $p < 0.001$), but no significant changes in Syrphidae ($LR\ \chi^2 = 0.22$, $df = 1$, $p = 0.64$) (Fig. 2D, Table 1, Suppl. material 4).

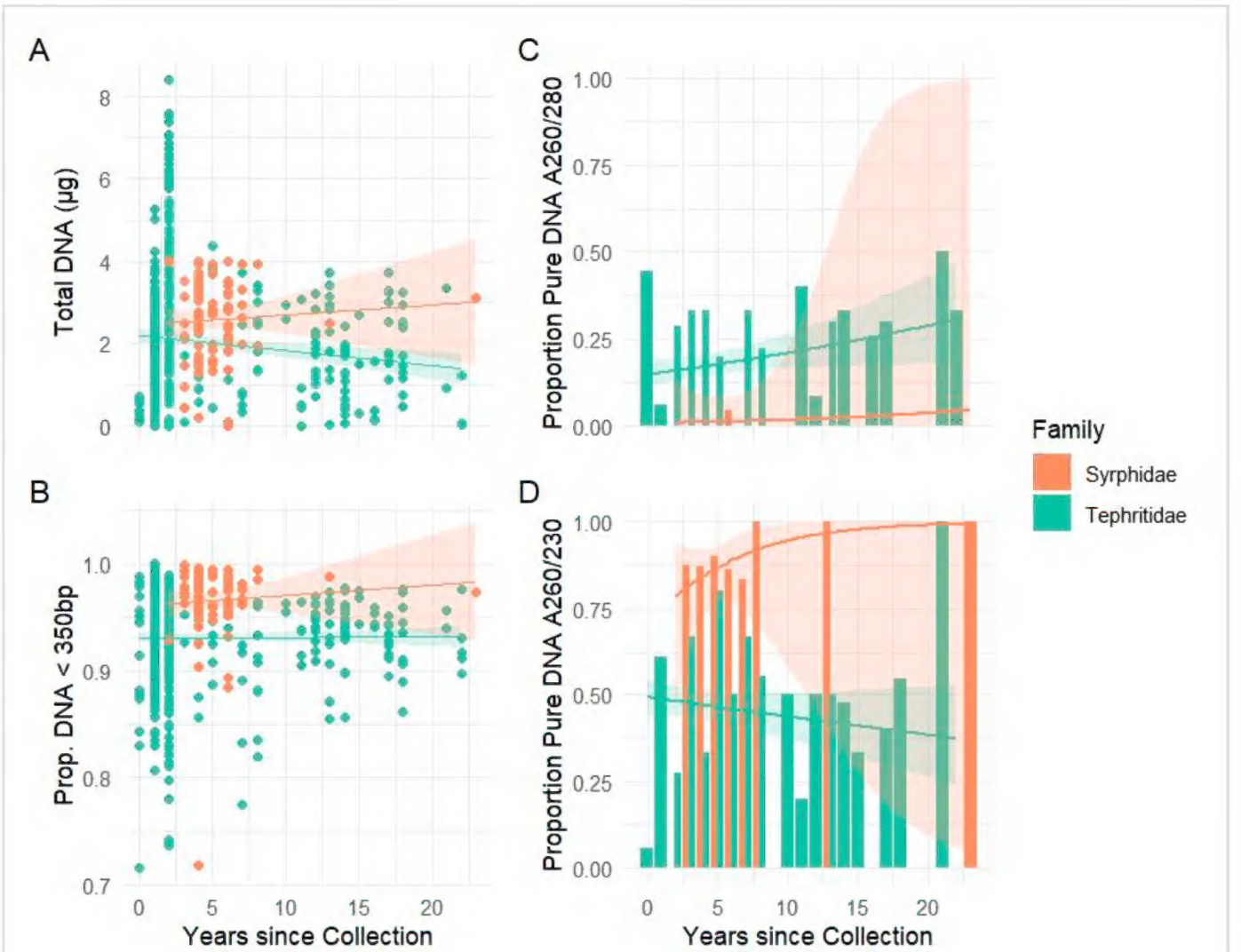

Figure 2. doi

DNA quantity and quality (linear regression lines are indicated for Syrphidae and Tephritidae). Relationships between years since collection and (**A**) total DNA recovered per specimen (µg), (**B**) proportion of the DNA molarity of short DNA fragments (< 350 bp), (**C**) proportion of “pure” DNA of A260/280 and (**D**) proportion of “pure” DNA of A260/230.

Out of 1,405 DNA extracts, 1,315 (93.6%) contained measurable amounts of DNA and were subjected to genomic library preparation and sequencing. Of these, 1,305 samples (99.2%) were successfully sequenced (i.e. provided WGS data of approximately 10x coverage). Overall, 92.9% of all samples processed in this study were successfully sequenced and produced genomic data for downstream genomic analyses, including population genomics, with some of these samples also contributing to the study by Deschepper et al. (2024a). The average cost for sequencing these samples, including

those that failed and those for which sequencing was outsourced to two companies, corresponded to €77.25 ± €39.01. This average reflects two pricing schemes: the majority of samples were sequenced at €73 per sample by one provider, while a subset was processed by a second provider at €128 per sample. For particularly low-input DNA extracts requiring specialised library preparation, the cost increased to €148 per sample. In addition to sequencing, the cost of DNA extraction using the DNeasy Blood & Tissue Kit (QIAGEN) was €5.70 per sample (excluding personnel costs).

The proportion of high-quality reads (i.e. with Phred quality score Q > 30) ranged from 0.86 to 0.96 in 1,305 samples, with an average of 0.92 (± 0.15). Voucher age significantly affected the proportion of high-quality reads in Tephritidae, with newer specimens showing a lower proportion ($F_{1,1078}$ = 11.98, p < 0.001). No significant effect of age was found in Syrphidae ($F_{1,88}$ = 0.53, p = 0.47) (Fig. 3A; Table 1; Suppl. material 4). The proportion of reads aligned to the *D. melanogaster* reference genome ranged from 0.0025 to 0.28, with an average of 0.10 (± 0.047). In Tephritidae, this proportion decreased significantly in more recent vouchers ($F_{1,106}$ = 38.80, p < 0.001), while no significant differences were found in Syrphidae ($F_{1,26}$ = 0.87, p = 0.36) (Fig. 3B; Table 1, Suppl. material 4).

GenTegra DNA 2D barcoded vials with a silica matrix (https://gentegra.com/gentegra-dna-2/) were used for long-term archiving of the extracted DNA at room temperature. The tube barcode and position were linked to the DaRWIN collection voucher (Adam et al. 2019). A numeric lab book is now available to collect specimen and laboratory information and automate data export to DaRWIN, via Visual Basic macros. An analysis study is ongoing to link DaRWIN with the external Laboratory Information Management System. In addition, the digital images produced for a subset of samples are available as links to the DaRWIN collection vouchers. The backup and long-term storage of these data was ensured by the network-attached storage systems of RMCA.

Discussion

This study focused on developing standardised wet-laboratory and bioinformatic pipelines for the efficient generation, preservation and integration of genomic data from the RMCA Diptera collections, with particular attention to their long-term incorporation into the collection management system. By anchoring genomic data to physical vouchers and curatorial metadata, we aimed to enhance the accessibility of museum-based genomic resources for scientific research, including applications in integrative taxonomy and insect systematics.

Balancing cost and uncertainty in WGS for insect museum collections

One of the main constraints limiting the routine production of genomic data from museum collections has been the cost associated with library preparation and sequencing, especially when working with large numbers of specimens. For this reason, partial genome sequencing strategies, such as reduced representation libraries (Ewart et al.

2019) or mitochondrial genomics (Timmermans et al. 2015), have traditionally been preferred for large-scale applications. We opted for standardised, commercially available and cost-effective lab pipelines instead of specialised protocols for ancient DNA or museomics (Straube et al. 2021, Andreeva et al. 2022), as these are not designed for routine HTS despite their potential of producing better results (Gauthier et al. 2020, Orlando et al. 2021, Hawkins et al. 2022). Rapid advances in sequencing technologies are steadily driving down costs and, as a result, routine and large-scale WGS of voucher specimens is becoming an increasingly viable and valuable approach for unlocking the genomic potential of entire museum collections (Gillett et al. 2014, Malakasi et al. 2019, Strijk et al. 2020, Mullin et al. 2022).

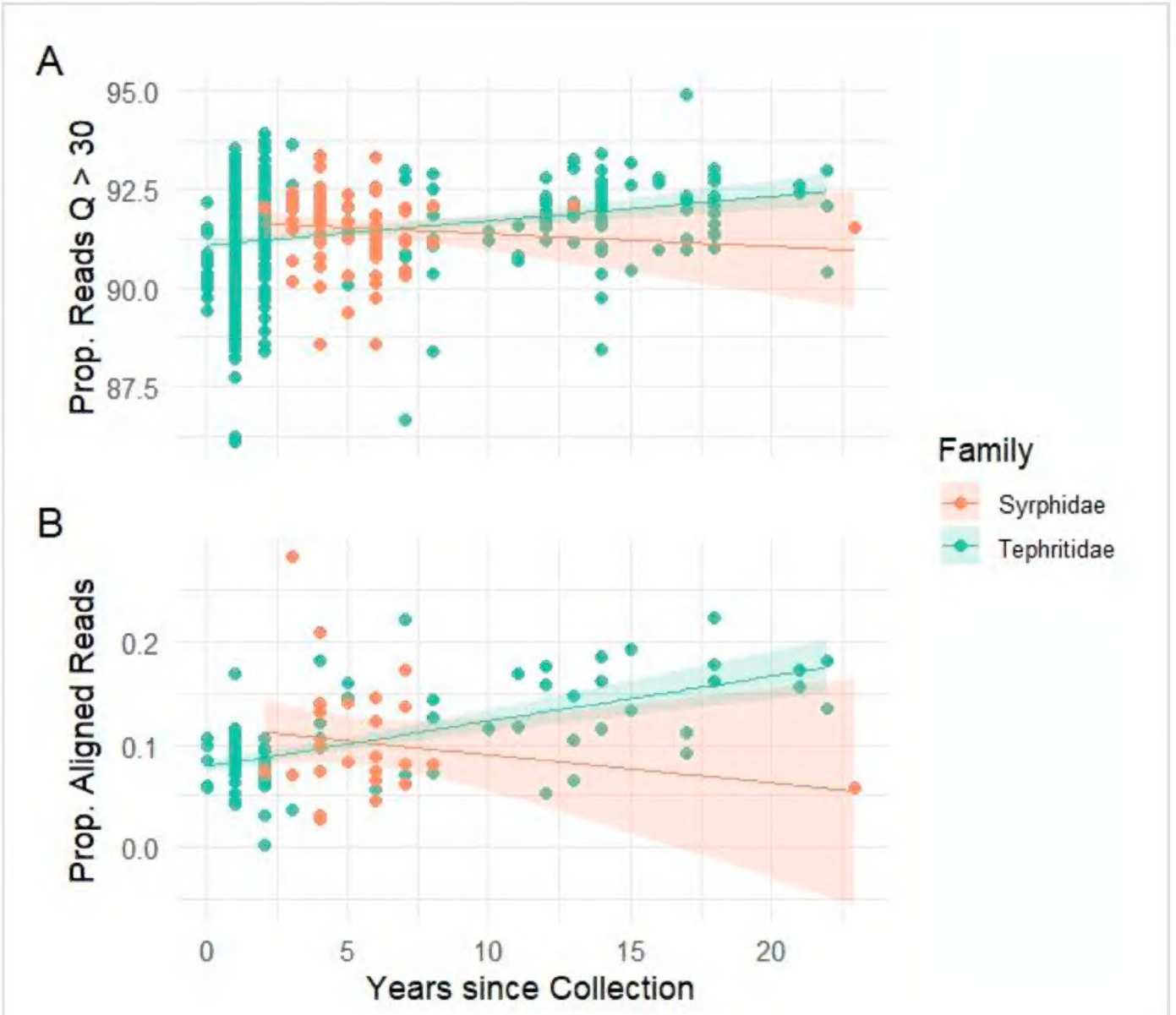

Figure 3. doi

Performance of whole genome sequencing (linear regression lines are indicated for Syrphidae and Tephritidae). Relationships between years since collection and (**A**) proportion of high-quality reads (Q > 30) and (**B**) proportion of reads aligned to the reference genome of *Drosophila melanogaster* (GCA_029775095.1).

Insect museum collections are often characterised by unbalanced sampling across taxa, time periods and preservation methods. In our study, this imbalance is compounded by largely unknown field sampling conditions for most specimens. These uncertainties

hinder our ability to discern clear patterns of sample suitability for genomics and the difficulty of predicting DNA quality or sequencing success is likely endemic to insect museum collections. Such high variability typically warrants a cautious approach, involving an in-depth assessment of each sample's DNA quality before applying targeted methods to resolve underlying degradation. However, the results of this study might offer a complementary perspective. Surprisingly and despite the relatively high proportion of suboptimal samples with highly fragmented or contaminated DNA, about 93% of our specimens were successfully sequenced using standard, cost- and time-effective protocols for DNA extraction, genomic library preparation and WGS. While we do not advocate indiscriminate sequencing of insect museum collections, our results show that Illumina short read WGS can work well even with degraded insect material. In this context, routine short-read sequencing offers a practical first step for genomic data mining, particularly for large collections. It enables us to reserve more complex and resource-intensive methods for the subset of samples that fail initial sequencing (which in our specific case represented 7% of specimens). This strategy could become increasingly viable as the costs of -omics technologies continue to decline. Its effectiveness may also improve as a growing number of high quality insect reference genomes (Deschepper et al. 2024b) allow for more accurate mapping and assembly of Illumina short read data.

DNA quality prediction hindered by sparse sampling history and storage conditions

The parameters considered in this feasibility study for the Illumina short-read WGS of the Diptera collections of RMCA revealed marked differences between the two insect families. Syrphidae samples demonstrated greater stability in DNA quality and sequencing metrics over time. Specifically, voucher age did not significantly affect DNA yield, DNA fragmentation (proportion of short DNA fragments), DNA quality or sequencing success (proportion of high-quality reads and alignment rates) in Syrphidae specimens, indicating relatively consistent preservation quality across the sampling window. In contrast, the DNA quantity of the Tephritidae specimens increased significantly in more recently collected vouchers. However, DNA quality metrics in Tephritidae showed clear age-related declines: neither the proportion of pure DNA samples, based on absorbance ratios, nor the proportion of high-quality sequencing reads improved with age. Instead, more recent Tephritidae samples exhibited higher DNA fragmentation and a significant decrease in the proportion of high-quality reads and aligned sequences, suggesting increased degradation or other factors negatively affecting sequencing performance in younger specimens. These differences likely reflect the uneven sampling efforts carried out over the years across the two collections, underscoring the complex interplay of biological and environmental processes that influence the suitability of insect specimens for genomic research. Previous studies have shown that specimen size, tissue type and preservation methods significantly affect DNA quality and yield (Kistler et al. 2017, Brewer et al. 2019, Card et al. 2021, Andreeva et al. 2022, Mullin et al. 2022). Additionally, differences between specimens, taxon groups or museum collections are likely influenced by intrinsic features of each insect group as well

as by the heterogeneous sampling, collection management or laboratory procedures implemented by different curators, taxonomists and molecular biologists employed across institutions.

In this articulated scenario, part of the differences in DNA yield, quality and contamination between Syrphidae and Tephritidae may be attributed to how the specimens were collected and stored in the field. As is often the case with specimens from natural history collections, detailed information on collection and preservation procedures in the field is lacking for most of our samples, limiting our ability to systematically assess their impact. We do know, however, that nearly all Tephritidae and Syrphidae specimens were preserved in ethanol hours, days or even weeks after being trapped. Although detailed metadata are missing, general differences in field collection and preservation practices between the two groups are known and provide useful context for interpreting the results. While most Syrphidae were collected via hand netting and immediately preserved in absolute ethanol, Tephritidae were primarily collected using McPhail traps. These traps, specifically developed for monitoring true fruit flies in area-wide surveillance programmes, sterile insect technique operations, orchards and quarantine zones (FAO and IAEA 2018), contain an insecticide and a chemical lure to attract the target insects. They are typically left in the field for periods ranging from hours to several weeks before being emptied, depending on logistical constraints and the research objectives. Hence, if the traps are surveyed after long periods, the DNA of trapped specimens is more exposed to degrading environmental conditions, especially under (sub)tropical conditions (Fowler et al. 2024). Controlled laboratory trials and field evaluations conducted by Fowler et al. (2024), showed that high ambient temperatures and relative humidity significantly accelerate DNA degradation in tephritid flies, especially when exposure times are prolonged and specimens are not immediately preserved. These findings are consistent with previous studies reporting that elevated temperature and humidity can have detrimental effects on DNA quality in insects (Mandrioli 2008, Zimmermann et al. 2008), especially when collection protocols do not involve direct preservation in ethanol (Ballare et al. 2019). Similar concerns have been raised in recent field evaluations, which showed that the duration and type of trap exposure can significantly influence the suitability of Tephritidae for downstream molecular diagnostics (Brown et al. 2025). This aligns with broader evidence that the persistence of amplifiable DNA is highly sensitive to environmental conditions, particularly high humidity and rainfall, which accelerate degradation over time when specimens are not promptly preserved (Lee et al. 2019). Prolonged exposure under such conditions has repeatedly been shown to reduce DNA quality, as demonstrated in multiple studies highlighting the combined effects of temperature, humidity and exposure duration on DNA integrity (Lund and Dissing 2004, Stevens et al. 2011, Martoni et al. 2021, Butterwort et al. 2022). In addition, the temporal patterns identified in Tephritidae likely reflect heterogeneity arising from the distinct histories of specific sample groups. A fairly large group of Tephritidae collected in 2020 and 2021 was part of a single research project and yielded particularly degraded specimens (Fig. 1). Although these samples were collected recently, the lack of detailed information limits our understanding of their precise handling and storage conditions in the field. Based on what is known about typical collection practices in our

research group, we can speculate that these specimens may have been subjected to prolonged, suboptimal conditions such as high temperature and humidity. This emphasises the critical role of field sampling protocols, including frequent servicing of traps and immediate storage in absolute ethanol. In this context, even limited metadata, such as the presence of a collection date or date range could help infer the likely collection method and exposure time, offering a valuable proxy for assessing preservation quality. More importantly, it highlights the challenges of reconstructing the collection history of samples collected under diverse field conditions that are often undocumented and beyond the curator's control. Therefore, the inherently heterogeneous nature of natural history collections makes it challenging to predict consistent patterns in DNA yield, quality or sequencing performance across samples. As a result, the common expectation that more recent samples will consistently provide better genomic data may often be unmet.

Integrated curation of genomic data

This pilot allowed integrating genomic data into DaRWIN, the open-source database platform developed by and currently in use at the Royal Belgian Institute of Natural Sciences (RBINS, https://darwin.naturalsciences.be/) and the Royal Museum for Central Africa (RMCA, https://darwinweb.africamuseum.be/). DaRWIN is built on PostgreSQL for data management and uses Symfony, a PHP framework, for its user interface and search engine. The platform's source code is openly available on GitHub (https://github.com/naturalsciences/Darwin). The archiving system offers a public search engine and a back-end for scientific and curatorial collection management. The system represents taxonomic hierarchies with synonyms and allows taxonomic types to be flagged in the digital collection. It also enables the management of loans and the automatic labelling of physical specimens. It offers an importation pipeline for batches of data from tab-delimited or spreadsheet documents that checks and validates the taxonomic hierarchy of imported data and guarantees the alignment of the new data with its normalised relational model (taxa, collectors and collection locations).

RMCA and RBINS have adopted a common archiving technology, while maintaining separate collection management systems to support institution-specific workflows. The adoption of a shared archiving platform enhances compatibility and coordination across research institutes, given the close collaboration between RMCA and RBINS, two Belgian Federal Scientific Institutions supported by the same funding agency (the Belgian Science Policy), sharing laboratory infrastructure as well as a Joint Experimental Molecular Unit (JEMU; https://www.jemu.be/).

As a result of this study, the DaRWIN interface has been enhanced to integrate links to genomic data and metadata alongside traditional collection information. Specifically, the platform can now track the full chain of custody by linking each source specimen to its corresponding tissue sample and DNA voucher. These records are also connected to external resources, including genomic datasets archived on the RMCA servers. Metadata from these records are publicly accessible on this DaRWIN page. Genomic data are

gradually made open access, for example, by providing GenBank accessions, following the publication of results from various research projects. Within the broader field of museomics, our work illustrates how a collection management system can be upgraded to support the archiving of genomic data and metadata. Treating genomic data not as ancillary, but as an integral component of collection management bridges advanced genomics and traditional curation of museum collections. Importantly, it also enhances the traceability of the genetic resources originating from natural history collections. By recording and publishing the full chain of custody from specimen to sequence, it promotes compliance with international agreements on sample properties and loans. In this context, we believe that the routine archiving and curation of genomic data would favour more reproducible, transparent and inclusive collection-based research.

In summary, our study shows that Illumina short-read WGS is a workable approach for large-scale sequencing of insect museum specimens using standardised and cost-effective protocols. We identified significant predictors of sequencing success, including taxon-specific effects and field collection methods, though their impact varied. Finally, by integrating genomic metadata into a curated database platform, we laid the groundwork for genomic traceability in museum collections. These results offer practical insights for future museomic initiatives and underscore the importance of metadata standardisation and curation in genomic research.

Acknowledgements

This study, coordinated by the Joint Experimental Molecular Unit (JEMU) of the Royal Belgian Institute for Natural Sciences (RBINS) and the AfricaMuseum (RMCA), was funded by InsectMOoD (BRAIN-BE 2.0, B2/202/P2/InsectMOoD, 2018 -2023). Synergies with projects DISPEST, AGROVEG and DIPODIP (financed by the Directorate-general Development Cooperation and Humanitarian Aid through a framework agreement with the RMCA), FFI-PM (EU, H2020, grant 818184), REACT (EU, H2020, grant 101059523) and SYNTHESYS+ (EU, H2020, grant 823827) and ISeBAF (BELSPO, BRAIN-be 2.0, B2/191/P1) allowed processing the samples from the Tephritidae and Syrphidae collections of RMCA. We thank Katerina Guschanski (Uppsala University, Sweden), Peter Hollingsworth (Royal Botanic Garden Edinburgh, UK), Francesca Scolari (National Research Council, Pavia, Italy) for sharing their valuable expertise and for critical and positive input to the InsectMOoD project. We are grateful to Mark Schutze, Daniele Sommaggio, Thierry Backeljau and Gontran Sonet for their constructive feedback on this manuscript. In the framework of the *decolonisation* and *inclusion* policies of the RMCA, our study brings together a balanced representation of authors from Europe and Africa. The intellectual and physical property of vouchers and genomic data used in this study are regulated by the Nagoya Protocol on Access and Benefit-sharing or, when this was not possible, by Mutually Agreed Terms (MATs) on the use of genetic resources which are inspired by the principles of the Nagoya Protocol. All documents on samples' intellectual and physical property are available upon request to the authors.

Funding program

BELSPO, BRAIN-BE 2.0: B2/202/P2/InsectMOoD and B2/191/P1/ISeBaf

Directorate-general Development Cooperation and Humanitarian Aid through a framework agreement with the RMCA: DISPEST, AGROVEG and DIPODIP

EU, H2020: FFI-PM (grant 818184), REACT (grant 101059523) and SYNTHESYS+ (grant 823827)

Conflicts of interest

The authors have declared that no competing interests exist.

References

- Adam M, Theeten F, Herpers J, Vandenberghe T, Semal P, Van den Spiegel D, Duchesne P (2019) DaRWIN: An open source natural history collections data management system. Biodiversity Information Science and Standards 3 https://doi.org/10.3897/biss.3.39054
- Andreeva TV, Malyarchuk AB, Soshkina AD, Dudko NA, Plotnikova MY, Rogaev EI (2022) Methodologies for ancient DNA extraction from bones for genomic analysis: approaches and guidelines. Russian Journal of Genetics 58 (9): 1017-1035. https://doi.org/10.1134/s1022795422090034
- Ballare K, Pope N, Castilla A, Cusser S, Metz R, Jha S (2019) Utilizing field collected insects for next generation sequencing: effects of sampling, storage, and DNA extraction methods. Ecology and Evolution 9 (24): 13690-705. https://doi.org/10.1002/ECE3.5756.
- Brecko J, Mathys A, Dekoninck W, Leponce M, VandenSpiegel D, Semal P (2014) Focus stacking: comparing commercial top-end set-ups with a semi-automatic low budget approach. A possible solution for mass digitization of type specimens. ZooKeys 464: 1-23. https://doi.org/10.3897/zookeys.464.8615
- Brewer GE, Clarkson JJ, Maurin O, Zuntini AR, Barber V, Bellot S, Biggs N (2019) Factors affecting targeted sequencing of 353 nuclear genes from herbarium specimens spanning the diversity of angiosperms. Frontiers in Plant Science 10 https://doi.org/10.3389/FPLS.2019.01102/BIBTEX
- Brown GW, Starkie ML, Fowler EV, Blacket MJ, Royer JE, Mayer DG, Souza NM, Cheesman J, Missenden B, Irvine M, Schutze MK (2025) Field assessment of current and improved surveillance traps for fruit flies (Diptera: Tephritidae) in Australia. Journal of Economic Entomology https://doi.org/10.1093/jee/toaf085
- Butterwort V, Dansby H, Zink FA, Tembrock LR, Gilligan TM, Godoy A, Braswell WE, Kawahara AY (2022) A DNA extraction method for insects from sticky traps: targeting a low abundance pest, *Phthorimaea absoluta* (Lepidoptera: Gelechiidae), in mixed species communities. Journal of Economic Entomology 115 (3): 844-851. https://doi.org/10.1093/jee/toac046

- Card DC, Shapiro B, Giribet G, Moritz C, Edwards SV (2021) Museum genomics. Annual Review of Genetics 55: 633-659. https://doi.org/10.1146/ANNUREV-GENET-071719-020506/1
- Carter D, Walker A (1999) Care and conservation of natural history collections. Chapter 7: 139-151.
- Chen S, Zhou Y, Chen Y, Gu J (2018) Fastp:an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34 (17): 884-890. https://doi.org/10.1093/BIOINFORMATICS/BTY560
- Colella J, Tigano A, MacManes M (2020) A linked-read approach to museomics: higher quality de novo genome assemblies from degraded tissues. Molecular Ecology Resources 20 (4): 856-70. https://doi.org/10.1111/1755-0998.13155
- Delatte H, Meyer M, Virgilio M (2019) Genetic structure and range expansion of *Zeugodacus cucurbitae* (Diptera: Tephritidae) in Africa. Bulletin of Entomological Research 109 (6): 713-22. https://doi.org/10.1017/S0007485319000026
- De Meyer M, Mwatawala M, Copeland R, Virgilio M (2016) Description of new *Ceratitis* species (Diptera: Tephritidae) from Africa, or how morphological and DNA data are complementary in discovering unknown species and matching sexes. European Journal of Taxonomy 233 https://doi.org/10.5852/ejt.2016.233
- De Meyer M, Goergen G, Jordaens K (2020a) Taxonomic revision of the Afrotropical hover fly genus *Senaspis* Macquart (Diptera, Syrphidae). ZooKeys 1003: 83-160. https://doi.org/10.3897/zookeys.1003.56557
- De Meyer M, Goergen G, Jordaens K (2020b) Taxonomic revision of the Afrotropical *Phytomia* Guérin-Méneville (Diptera: Syrphidae). Zootaxa 4803 (2). https://doi.org/10.11646/zootaxa.4803.2.1
- De Meyer M, Goergen G, Midgley J, Jordaens K (2024) On the identity of the Afrotropical species of *Mallota* Meigen (Diptera: Syrphidae). European Journal of Taxonomy 958 https://doi.org/10.5852/ejt.2024.958.2675
- Deschepper P, Vanbergen S, Virgilio M, Sciarretta A, Colacci M, Rodovitis V, Jaques J, Bjeliš M, Bourtzis K, Papadopoulos N, De Meyer M (2024a) Global invasion history with climate-related allele frequency shifts in the invasive Mediterranean fruit fly (Diptera, Tephritidae: *Ceratitis capitata*). Scientific Reports 14 (1). https://doi.org/10.1038/s41598-024-76390-1
- Deschepper P, Vanbergen S, Esselens L, Terblanche J, Karsten M, Snyman M, Cugala D, Canhanga L, Bota L, Mwatawala M, Ramadhani M, Kudra A, Tairo J, Bakengesa J, Addison P, Manrakhan A, Gledel C, Delatte H, De Meyer M, Virgilio M (2024b) A new genome sequence resource for five invasive fruit flies of agricultural concern: *Ceratitis capitata*, *C. quilicii*, *C. rosa*, *Zeugodacus cucurbitae* and *Bactrocera zonata* (Diptera, Tephritidae). F1000Research 13 https://doi.org/10.12688/f1000research.157946.1
- Ewart K, Johnson R, Ogden R, Joseph L, Frankham G, Lo N (2019) Museum specimens provide reliable SNP data for population genomic analysis of a widely distributed but threatened cockatoo species. Molecular Ecology Resources 19 (6): 1578-1592. https://doi.org/10.1111/1755-0998.13082
- FAO, IAEA (2018) Trapping systems. In: Enkerlin WR, Reyes-Flores J (Eds) Trapping guidelines for area-wide fruit fly programmes. 2nd ed.
- Ferrari G, Esselens L, Hart M, Janssens S, Kidner C, Mascarello M, Peñalba J, Pezzini F, von Rintelen T, Sonet G, Vangestel C, Virgilio M, Hollingsworth P (2023) Developing the protocol infrastructure for DNA sequencing natural history collections. Biodiversity Data Journal 11 https://doi.org/10.3897/bdj.11.e102317

- Fowler E, Starkie M, Blacket M, Mayer D, Schutze M (2024) Effect of temperature and humidity on insect DNA integrity evaluated by real-time PCR. Journal of Economic Entomology 117 (5): 1995-2002. https://doi.org/10.1093/JEE/TOAE193.
- Gauthier J, Pajkovic M, Neuenschwander S, Kaila L, Schmid S, Orlando L, Alvarez N (2020) Museomics identifies genetic erosion in two butterfly species across the 20th century in Finland. Molecular Ecology Resources 20 (5): 1191-1205. https://doi.org/10.1111/1755-0998.13167
- Gillett CD, Crampton-Platt A, Timmermans MT, Jordal B, Emerson B, Vogler A (2014) Bulk de novo mitogenome assembly from pooled total DNA elucidates the phylogeny of weevils (Coleoptera: Curculionoidea). Molecular Biology and Evolution 31 (8): 2223-2237. https://doi.org/10.1093/molbev/msu154
- Guschanski K, Krause J, Sawyer S, Valente L, Bailey S, Finstermeier K, Sabin R, Gilissen E, Sonet G, Nagy Z, Lenglet G, Mayer F, Savolainen V (2013) Next-generation museomics disentangles one of the largest primate radiations. Systematic Biology 62 (4): 539-554. https://doi.org/10.1093/sysbio/syt018
- Hawkins MR, Flores M, McGowen M, Hinckley A (2022) A comparative analysis of extraction protocol performance on degraded mammalian museum specimens. Frontiers in Ecology and Evolution 10 (August). https://doi.org/10.3389/FEVO.2022.984056/BIBTEX
- Kistler L, Ware R, Smith O, Collins M, Allaby R (2017) A new model for ancient DNA decay based on paleogenomic meta-analysis. Nucleic Acids Research 45 (11): 6310-20. https://doi.org/10.1093/NAR/GKX361
- Knyshov A, Gordon EL, Weirauch C (2019a) Cost-efficient high throughput capture of museum arthropod specimen DNA using PCR -generated baits. Methods in Ecology and Evolution 10 (6): 841-52. https://doi.org/10.1111/2041-210X.13169.
- Knyshov A, Hoey-Chamberlain R, Weirauch C (2019b) Hybrid enrichment of poorly preserved museum specimens refines homology hypotheses in a group of minute litter bugs (Hemiptera: Dipsocoromorpha: Schizopteridae). Systematic Entomology 44 (4): 985-95. https://doi.org/10.1111/syen.12368.
- Lee LYC, Wong HY, Lee JY, Waffa ZBM, Aw ZQ, Fauzi SNABM, Hoe SY, Lim ML, Syn CKC (2019) Persistence of DNA in the Singapore Context. International Journal of Legal Medicine 133 (5): 1341-49. https://doi.org/10.1007/S00414-019-02077-2/METRICS.
- Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25 (14): 1754-60. https://doi.org/10.1093/BIOINFORMATICS/BTP324
- Lucena-Aguilar G, Sánchez-López AM, Barberán-Aceituno C, Carrillo-Ávila JA, López-Guerrero JA, Aguilar-Quesada R (2016) DNA source selection for downstream applications based on DNA quality indicators analysis. Biopreservation and Biobanking 14: 264-70. https://doi.org/10.1089/bio.2015.0064
- Lund S, Dissing J (2004) Surprising stability of DNA in stains at extreme humidity and temperature. International Congress Series 1261.
- Malakasi P, Bellot S, Dee R, Grace O (2019) Museomics clarifies the classification of *Aloidendron* (Asphodelaceae), the iconic African tree aloes. Frontiers in Plant Science 10 https://doi.org/10.3389/fpls.2019.01227
- Mandrioli M (2008) Insect collections and DNA analyses: how to manage collections? Museum Management and Curatorship 23 (2): 193-99. https://doi.org/10.1080/09647770802012375.

- Martoni F, Nogarotto E, Piper A, Mann R, Valenzuela I, Eow L, Rako L, Rodoni B, Blacket M (2021) Propylene glycol and non-destructive DNA extractions enable preservation and isolation of insect and hosted bacterial DNA. Agriculture 11 (1): 77. https://doi.org/10.3390/agriculture11010077
- Mullin V, Stephen W, Arce A, Nash W, Raine C, Notton D, Whiffin A, Blagderov V, Gharbi K, Hogan J, Hunter T, Irish N, Jackson S, Judd S, Watkins C, Haerty W, Ollerton J, Brace S, Gill R, Barnes I (2022) First large-scale quantification study of DNA preservation in insects from natural history collections using genome-wide sequencing. Methods in Ecology and Evolution 14 (2): 360-371. https://doi.org/10.1111/2041-210x.13945
- Nakahama N (2020) Museum specimens: an overlooked and valuable material for conservation genetics. Ecological Research 36 (1): 13-23. https://doi.org/10.1111/1440-1703.12181
- Orlando L, Allaby R, Skoglund P, Der Sarkissian C, Stockhammer P, Ávila-Arcos M, Fu Q, Krause J, Willerslev E, Stone A, Warinner C (2021) Ancient DNA analysis. Nature reviews methods primers 1 (1). https://doi.org/10.1038/s43586-020-00011-0
- Secretariat Biological Diversity (2011) Nagoya protocol on access to genetic resources and the fair and equitable sharing of benefits arising from their utilization to the convention on biological diversity. United Nations URL: https://www.cbd.int/abs
- Springer Nature (2022) Nature addresses helicopter research and ethics dumping. Nature 606 (7912): 7-7. https://doi.org/10.1038/d41586-022-01423-6
- Stevens M, Warren G, Mo J, Schlipalius D (2011) Maintaining DNA quality in stored-grain beetles caught in lindgren funnel traps. Journal of Stored Products Research 47 (2): 69-75. https://doi.org/10.1016/J.JSPR.2010.10.002.
- Straube N, Lyra M, Paijmans JA, Preick M, Basler N, Penner J, Rödel M, Westbury M, Haddad CB, Barlow A, Hofreiter M (2021) Successful application of ancient DNA extraction and library construction protocols to museum wet collection specimens. Molecular Ecology Resources 21 (7): 2299-2315. https://doi.org/10.1111/1755-0998.13433
- Strijk J, Binh HT, Ngoc NV, Pereira J, Slik JWF, Sukri R, Suyama Y, Tagane S, Wieringa J, Yahara T, Hinsinger D (2020) Museomics for reconstructing historical floristic exchanges: Divergence of stone oaks across Wallacea. PLOS One 15 (5). https://doi.org/10.1371/journal.pone.0232936
- Timmermans MTN, Viberg C, Martin G, Hopkins K, Vogler A (2015) Rapid assembly of taxonomically validated mitochondrial genomes from historical insect collections. Biological Journal of the Linnean Society 117 (1): 83-95. https://doi.org/10.1111/bij.12552
- Zimmermann J, Hajibabaei M, Blackburn D, Hanken J, Cantin E, Posfai J, Evans T (2008) DNA damage in preserved specimens and tissue samples: a molecular sssessment. Frontiers in Zoology 5 (1): 1-13. https://doi.org/10.1186/1742-9994-5-18/FIGURES/9.
- Zuur AF, Hilbe JM, Ieno EN (2013) A beginner's guide to GLM and GLMM with R: a frequentist and Bayesian perspective for ecologists. Highland Statistics Ltd.

Supplementary materials

Suppl. material 1: Species list doi

Authors: Lore Esselens
Data type: Table
Brief description: Species list of specimens processed in this study.
Download file (150.60 kb)

Suppl. material 2: DNA extraction kits doi

Authors: Lore Esselens
Data type: Table
Brief description: DNA extraction kits used for the preliminary comparisons.
Download file (62.54 kb)

Suppl. material 3: DNA yields from DNA extraction protocols doi

Authors: Lore Esselens
Data type: Graph
Brief description: Total DNA yields from four DNA extraction protocols.
Download file (31.04 kb)

Suppl. material 4: ANCOVA doi

Authors: Lore Esselens
Data type: Table
Brief description: Analysis of Covariance of the general(ised) linear models testing.
Download file (186.32 kb)